REPORT DOCUMENTATION PAGE	*Form Approved* *OMB No. 0704-0188*

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
03-01-2006	Final Report	04/30/00 - 12/31/05

4. TITLE AND SUBTITLE
Strategic and Tactical Decision-Making Under Uncertainty

5a. CONTRACT NUMBER N00014-00-1-0637
5b. GRANT NUMBER n/a
5c. PROGRAM ELEMENT NUMBER n/a
5d. PROJECT NUMBER n/a
5e. TASK NUMBER n/a
5f. WORK UNIT NUMBER n/a

6. AUTHOR(S)
Jordan, Michael I.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
The Regents of the University of California
Sponsored Projects Office
2150 Shattuck Avenue, Suite 313
Berkeley, CA 94720-5940

8. PERFORMING ORGANIZATION REPORT NUMBER
012243-009

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Wendy L. Martinez
Office of Naval Research
875 N. Randolph St.
One Liberty Center
Arlington, VA 22203-1995

10. SPONSOR/MONITOR'S ACRONYM(S)
ONR 311

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
n/a

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for Public Release; distribution is Unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT
This report presents the final conclusions of the research on decision-making under uncertainty conducted by the investigators at the University of California at Berkeley, Stanford University, and the University of California at Davis, under the aegis of the MURI on Decision-Making under Uncertainty.

15. SUBJECT TERMS
Markov decision processes, graphical models, reinforcement learning, statistical learning, decentralized control

16. SECURITY CLASSIFICATION OF: a. REPORT	b. ABSTRACT	c. THIS PAGE	17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
U	U	U	SAR	20	Michael I. Jordan

19b. TELEPHONE NUMBER *(Include area code)*
510/642-3806

Standard Form 298 (Rev. 8/98)
Prescribed by ANSI Std. Z39.18

FINAL REPORT

March 1, 2006

Strategic and Tactical Decision–Making under Uncertainty

Michael Jordan (PI), Venkat Anantharam, Laurent El Ghaoui,
Stuart Russell, and Shankar Sastry
University of California, Berkeley

Daphne Koller (co-PI), Benjamin Van Roy and Claire Tomlin
Stanford University

Roger Wets
University of California, Davis

This report presents the final conclusions of the research conducted by the investigators listed above at the University of California at Berkeley, Stanford University, and the University of California at Davis, under the aegis of the MURI on Decision-Making under Uncertainty. The MURI was a significant success, with its accomplishments far surpassing the goals of the original proposal. Even at this early date, it is clear that several of our research outcomes are being viewed as seminal contributions to the literature.

1 Highlights

While our detailed presentation in this report will focus principally on our work in the final period of the grant—augmenting the progress reports submitted in previous years—we begin by highlighting some of the major intellectual accomplishments of the MURI throughout the grant's lifetime. The work that we highlight here has been chosen in part because it is highly-cited (with in many cases hundreds of citations as measured by Google Scholar and Citeseer) and has led to substantial follow-up research and independent implementations and applications in various literatures. We also highlight some of the educational and tech transfer issues surrounding the project.

1.1 Major intellectual accomplishments

- *Factored Markov decision processes*: Through the development of factored Markov decision processes we have been able to solve sequential decision-making problems that are many orders of magnitude larger than those studied in previous work.
- *Decomposed decision-making in hierarchical systems*: We have developed new architectures for decomposed decision-making that involve an interaction between knowledge-based and learning-based formalisms within an overall "partial program + reinforcement learning" approach. This has included the development of ALISP, an extension of LISP that allows for partial specification of agent programs.

- *Linear programming approximations for Markov decision problems*: We have developed a broadly useful computational approach to sequential decision making under uncertainty based on a linear programming formulation for approximate dynamic programming. We have developed strong theory for this approach, and experiments with large-scale test beds involving queueing networks and server farm management have demonstrated the practical utility of the approach.
- *Policy gradient for reinforcement learning*: Our development of a new approach to reinforcement learning known as the "PEGASUS algorithm" has made it possible to solve challenging partially-observable MDP problems. In particular, PEGASUS has successfully flown the Berkeley autonomous helicopter and has been adopted for a variety of nonlinear adaptive control projects.
- *Pursuit-evasion games*: We have developed new algorithms for solving distributed pursuit-evasion games and have demonstrated the new capabilities that they offer for situation assessment and distributed control on our hardware platform. This has involved interacting UAVs and ground-based robots.
- *Equilibrium classification*: We have generated a unifying framework for equilibrium classification within the feasible set of solutions of centralized optimization problem using local Lagrange multipliers and tangent spaces. Such classifications lead to a particular type of solution such as feasible, Nash, first-order Pareto-optimal or second order Pareto-optimal solutions within the centralized optimization problem.
- *Large-margin Markov networks*: We have presented a widely-applicable solution strategy for the problem of discriminative training of graphical model and combinatorial learning systems. This strategy has been widely adopted in the literature.
- *Kernel dimensionality reduction (KDR)*: We developed a novel mathematical framework for representing conditional independence assertions using cross-covariance operators on reproducing kernel Hilbert spaces. This framework leads to the first fully nonparametric methodology for solving the sufficient dimension reduction problem in regression and classification problems.
- *Convex optimization algorithms for machine learning*: We have pioneered the use of tools from convex optimization theory in the field of machine learning, including a new minimax classification algorithm, a new class of techniques for learning kernel matrices based on semidefinite programming, and new algorithmic procedures based on extragradient and dual extragradient methods.
- *Latent Dirichlet allocation (LDA)*: LDA is a Bayesian latent variable model for representing data clusters, in which entities can belong to more than one cluster, and in which cluster membership probabilities are modeled explicitly. This model has been widely used in numerous literatures, including social networks, computational vision, information retrieval and computational linguistics.
- *Kernel independent component analysis (KICA)*: KICA is a novel approach to the independent component analysis problem that allows the power of kernel methods to be brought to bear on a semiparametric statistical modeling problem. KICA has become widely cited as a state-of-the-art algorithms for problems in source separation.

1.2 Education

The following is a list of the student and postdoctoral contributors to the MURI project, accompanied by their current position in academia or industry. The success of our students in landing positions at the most prominent academic and industrial enterprises is one of the major indications of the success of the MURI and may well be its most lasting legacy:

- Alexandre d'Aspremont, Assistant Professor, Princeton University
- Francis Bach, Assistant Professor, Ecole des Mines
- David Blei, Assistant Professor, Princeton University
- Eyal Amir, Assistant Professor, University of Illinois at Urbana-Champaign
- Michael Casey, Research Scientist, Raytheon
- Daniela Pucci de Farias, Assistant Professor, MIT
- Nando de Freitas, Assistant Professor, University of British Columbia
- Carlos Guestrin, Assistant Professor, Carnegie Mellon University
- Gokhan Inalhan, Postdoctoral Fellow, MIT
- Gert Lanckriet, Assistant Professor, University of California, San Diego
- Bhaskara Marthi, Postdoctoral Fellow, MIT
- Jon McAuliffe, Assistant Professor, University of Pennsylvania
- Brian Milch, Postdoctoral Fellow, MIT
- Kevin Murphy, Assistant Professor, University of British Columbia
- Andrew Ng, Assistant Professor, Stanford University
- Payam Pakzad, Postdoctoral Fellow, Ecole Polytechnique Federale de Lausanne
- Mark Paskin, Research Scientist, Google
- Hanna Pasula, Postdoctoral Fellow, MIT
- Paat Rusmevichientong, Assistant Professor, Cornell University
- Christian Shelton, University of California, Riverside
- Dusan Stipanovic, Assistant Professor, University of Illinois at Urbana-Champaign
- Ben Taskar, Assistant Professor, University of Pennsylvania
- Sekhar Tatikonda, Assistant Professor, Yale University
- Rene Vidal, Assistant Professor, Johns Hopkins University
- Martin Wainwright, Assistant Professor, University of California, Berkeley
- Eric Xing, Assistant Professor, Carnegie Mellon University
- Alice Zheng, Postdoctoral Fellow, Carnegie Mellon University

1.3 Industry and Government Collaboration

We have had extensive interactions over the course of the MURI project with groups in industry and government, interactions which have catalyzed the transfer of our research results into applied settings. These include:

- Alphatech (Bob Washburn, John Fox)
- Intel Corporation (Gary Bradski, John Marc Agosta)
- Nasa Ames (George Meyer, Asaf Degani, Len Tobias)
- Inxight (Ramana Rao)

- DARPA ITO (John Bay)
- AFOSR (Belinda King)
- AFRL (Siva Banda)
- Google (Peter Norvig)
- Defense Threat Reduction Agency
- Vismod (Doron Tal)

2 Brief Overview of Research

This section provides a brief overview of our main results in the final period of the grant. Fuller descriptions are provided in Section 3.

Markov Decision Algorithms

We developed a new algorithm based on linear programming for optimization of average-cost Markov decision processes (MDPs). The algorithm approximates the differential cost function of a perturbed MDP via a linear combination of basis functions.

We also obtained performance loss bounds for approximate value iteration with state aggregation and extended the result to a case which incorporates exploration.

Decentralized Estimation and Control

We have developed a new class of algorithms for solving the distributed detection problem. Rather than assuming that the class-conditional densities are known—an unrealistic assumption that has permeated the literature for two decades—our approach is based solely on empirical data.

We have developed a distributed game theory model, motivated by the possibility of using externally provided common randomness to improve the performance of distributed systems.

We have developed a control-theoretic formulation of the phenomenon of stochastic resonance, where noise enhances the ability of an observer to detect the presence of a weak perturbing signal in certain dynamical systems.

Multi-Agent Systems

We have shown that systems based on extracting eigenvectors are prone to undesirable gaming behavior by colluders. We developed an algorithm that detects such collusive behavior and can be used to produce a more robust version of PageRank. We provide theorems demonstrating the effectiveness of the method.

Graphical Model Representations

We have developed a formalism that allows modeling and reasoning for continuous time graphical models. This formalism yields answers to questions about when events will happen or when current or future conditions will stop. We have developed a framework and two associated algorithms for inference in these networks.

Graphical Model Inference

We have developed a systematic theory of approximate inference algorithms that encompasses all of the variational inference algorithms developed earlier in this MURI project.

In particular, we view approximate inference as a relaxation of a particular family of non-convex optimization problems to the optimization of convex approximations over a marginal polytope.

We have developed the minimal graphical representation theory for the Kikuchi approximation method for estimation in the context of a magnetic recording application. The method outperforms the currently used methods in the regime of most practical interest, and may therefore have an important impact.

Learning

We have developed a sparse form of PCA. In particular we have shown how to solve the problem of approximating, in the Frobenius-norm sense, a positive, semidefinite symmetric matrix by a rank-one matrix, with an upper bound on the cardinality of its eigenvector.

We have developed a simple and scalable algorithm for maximum-margin estimation of structured output models, including an important class of Markov networks and combinatorial models. We formulated the estimation problem as a convex-concave saddle-point problem that allows us to use simple projection methods based on the dual extragradient algorithm.

We developed a novel mathematical framework for representing conditional independence assertions using cross-covariance operators on reproducing kernel Hilbert spaces. This framework leads to the first fully nonparametric methodology for solving the sufficient dimension reduction problem in regression and classification.

3 Detailed Description of Research

We now give a detailed description of our major research results in the final period of the grant.

3.1 Markov Decision Algorithms

Linear Program for Bellman Error Minimization with Performance Guarantees

We introduced a new algorithm based on linear programming for optimization of average-cost Markov decision processes (MDPs). The algorithm approximates the differential cost function of a perturbed MDP via a linear combination of basis functions. The approximation minimizes a version of Bellman error. We establish an error bound that scales gracefully with the number of states without imposing the (strong) Lyapunov condition required by its counterpart in our prior work on linear programming methods for approximate dynamic programming.

Performance Loss Bounds for Approximate Value Iteration with State Aggregation

We studied solutions to an equation that characterizes fixed points of a certain form of value iteration or stationary points of "on-policy" temporal-difference learning when the value function is approximated via state aggregation (i.e., the state space is partitioned, and values within each partition are approximated by a constant). We established a performance bound showing that this solution provides close approximations that are superior to the standard form of approximate value iteration by an arbitrarily large factor. A solution to

the equation of interest does not always exist. We also studied a modified form of the equation which incorporates "exploration" and for which a solution is guaranteed to exist. We have established that the solution to this equation satisfies a similar performance bound.

3.2 Decentralized Estimation and Control

Decentralized Detection

We consider the problem of decentralized detection under constraints on the number of bits that can be transmitted by each sensor. In contrast to most previous work, in which the joint distribution of sensor observations is assumed to be known, we address the problem when only a set of empirical samples is available. We propose a novel algorithm using the framework of empirical risk minimization and marginalized kernels, and analyze its computational and statistical properties both theoretically and empirically. We provide an efficient implementation of the algorithm, and demonstrate its performance on both simulated and real data sets.

As part of this line of work, we have also derived a general theorem that establishes a correspondence between surrogate loss functions in classification and the family of f-divergences. Moreover, we have provided constructive procedures for determining the f-divergence induced by a given surrogate loss, and conversely for finding all surrogate loss functions that realize a given f-divergence. We introduced the notion of universal equivalence among loss functions and corresponding f-divergences, and provided necessary and sufficient conditions for universal equivalence to hold. These ideas have applications to classification problems that also involve a component of experiment design; in particular, we have leveraged our results to prove consistency of our decentralized detection procedure.

Distributed Game Theory Model

We have studied the use of externally provided randomness to improve the performance of distributed engineering systems. The possibility of doing this emerged in work on cryptography by Ueli Maurer in the early 90's. One basic idea is that a distributed system can create a larger family of joint distributions on its actions if the agents are provided with common randomness. We have developed this idea in the context of a novel distributed game theory formulation, which models one of the players of the game as being represented by a distributed agent. This formulation may be of particular interest in pursuit evasion scenarios where the evader is being pursued by distributed pursuers; for instance the "pursuers" may be the motes of a sensor network.

Stochastic Resonance

Externally provided randomness can enhance performance for the detection of weak signals through the phenomenon of "stochastic resonance." We have developed a control theory for the open loop choice of the noise to maximize stochastic resonance, where the efficiency of the resonance is measured in an information theoretic way.

3.3 Multi-Agent Systems

Making PageRank Robust to Collusion

We consider the problem of computing a robust Pagerank given only the link structure of the web, interpreting Pagerank as the stationary distribution of a Markov process. The robustness is with respect to collusion, whereby the link structure of some groups of nodes is such that their stationary distribution is higher than it "should" be. We first develop a mathematical definition of collusion that captures our intuitive understanding of the concept, noting that intent is undetectable as we are only given the link structure.

The mathematical definition is computationally impractical to compute, as it involves searching over all subsets of nodes. A computationally feasible method is developed, looking at the sensitivity of a sites Pagerank with respect to changes in reset parameters. Two variations are considered, and theorems are developed for both, demonstrating the effectiveness of the detection methods under certain conditions.

3.4 Graphical Model Representations

Continuous-Time Graphical Models

We have developed a formalism that allows modeling and reasoning for continuous time graphical models. This formalism yields answers to questions about when events will happen or when current or future conditions will stop. We have developed a framework and two associated algorithms for inference in these networks and have developed learning algorithms to extract continuous time Bayesian networks from data. We have also developed error bounds for continuous-time posterior probabilities under Poisson sampling of continuous-time models.

3.5 Graphical Model Inference

Variational Inference

We have developed a systematic theory of approximate inference algorithms that encompasses all of the variational inference algorithms developed earlier in this MURI project. In particular, we view approximate inference as a relaxation of a particular family of non-convex optimization problems to the optimization of convex approximations over a marginal polytope. Mean field algorithms are characterized as being an inner approximation to the marginal polytope, and belief propagation algorithms are characterized as being an outer approximation to the marginal polytope. This unification has also led to new algorithms based on semidefinite outer relaxations of the marginal polytope.

Kikuchi Approximation

Bayesian belief propagation is known to have connections with statistical mechanics. Specifically, a message passing algorithms can be viewed as an iterative update rule for the Lagrange multipliers in the problem of constrained minimization of a free energy functional associated to the local functions that define the estimation problem, the constraints being certain consistency conditions. This is the so-called Bethe approximation of statistical physics, which is a special case of a more general so called Kikuchi approximation. This led to the speculation that algorithms related to the updating of Lagrange multipliers in the Kikuchi approximation could give new efficient ways of solving estimation problems.

In an earlier report we presented our discovery that a certain kind of diamond structure

in the consistency constraints of the general Kikuchi approximation implies redundancies, can be consistently removed to reduce the complexity of the associated message passing algorithms. In recent work we applied this method to the problem of joint decoding of a low-density parity-check (LDPC) code and a partial response (PR) channel (this is the kind of channel relevant to magnetic recording applications). As we proved earlier, our approach of breaking diamonds results in the minimal graphical representation; algorithms on such a representation have the fewest messages and are hence the least complex per each iteration. In examples we studied, we found a significant improvement in decoding performance using our approach, relative to the currently used approaches, particularly at low SNR (signal to noise ratio), which is the regime of most interest for these applications. The price to pay for this is an increased cost in the number messages (by about 66% in typical examples). However, it is not possible to get improved performance by paying this cost using the traditional approach, so our ideas appear to be of importance.

3.6 Learning

Sparse Principal Component Analysis

Principal component analysis (PCA) is a popular tool for data analysis and dimensionality reduction. It has applications throughout science and engineering. In essence, PCA finds linear combinations of the variables (the so-called principal components) that correspond to directions of maximal variance in the data. It can be performed via a singular value decomposition (SVD) of the data matrix A, or via an eigenvalue decomposition if A is a covariance matrix. We have developed a sparse form of PCA. In particular we have shown how to solve the problem of approximating, in the Frobenius-norm sense, a positive, semidefinite symmetric matrix by a rank-one matrix, with an upper bound on the cardinality of its eigenvector. The problem has numerous applications in communications and control. We used a modification of the classical variational representation of the largest eigenvalue of a symmetric matrix, where cardinality is constrained, and we have shown how the problem can be solved via a semidefinite programming based relaxation.

Structured Prediction, Dual Extragradient and Bregman Projections

We have developed a simple and scalable algorithm for maximum-margin estimation of structured output models, including an important class of Markov networks and combinatorial models. We formulated the estimation problem as a convex-concave saddle-point problem that allows us to use simple projection methods based on the dual extragradient algorithm. The projection step can be solved using dynamic programming or combinatorial algorithms for min-cost convex flow, depending on the structure of the problem. We showed that this approach provides a memory-efficient alternative to formulations based on reductions to a quadratic program (QP). We analyzed the convergence of the method and presented experiments on two very different structured prediction tasks: 3D image segmentation and word alignment, illustrating the favorable scaling properties of our algorithm.

Kernel Dimension Reduction

We have proposed a novel method of dimensionality reduction for supervised learning problems. Given a regression or classification problem in which we wish to predict a re-

sponse variable Y from an explanatory variable X, we have shown how to treat the problem of dimensionality reduction as that of finding a low-dimensional "effective subspace" for X which retains the statistical relationship between X and Y. We have shown that this problem can be formulated in terms of conditional independence. To turn this formulation into an optimization problem we established a general nonparametric characterization of conditional independence using covariance operators on reproducing kernel Hilbert spaces. This characterization allowed us to derive a contrast function for estimation of the effective subspace. Unlike many conventional methods for dimensionality reduction in supervised learning, the proposed method requires neither assumptions on the marginal distribution of X, nor a parametric model of the conditional distribution of Y. We presented experiments that compare the performance of the method with conventional methods.

4 Testbeds and Software

4.1 Autonomous Helicopter and Robot Locomotion

The PEGASUS algorithm (Ng and Jordan, 2000) gives an efficient way of finding good control policies for very large POMDPs. The key in PEGASUS lies in its searching for policies by evaluating each of them on a small number of "representative scenarios." As a concrete example, if we are learning to quickly locate mines randomly buried in a road, a "scenario" might be a specific placement of the mines. As another example, if we are trying to track and intercept an unguided missile, the scenario might be the specific (random) trajectory taken by the missile. In many situations, it is not clear what a representative "scenario" is, but we show they can be automatically defined and generated for *any* POMDP. This stems from the rather surprising mathematical fact that any stochastic POMDP problem can be reduced to one with only deterministic transition dynamics. We also proved that in order to find good policies, the number of scenarios needed is small—generally only a low-order polynomial of the dimension of the state space. This should be contrasted with methods that discretize the state space or that attempt exact solutions to the Bellman equations, which suffer from the curse of dimensionality and hence are totally inapplicable to even moderate-sized problems.

We have successfully applied this method to the challenging problem of flying an autonomous helicopter. Helicopters have complex, non-linear, stochastic, and highly coupled dynamics, and present a high-dimensional, challenging, MIMO (multiple-input multiple-output) problem. On our first attempt, the PEGASUS-learned policy flew the helicopter significantly more stably than a human pilot. To date, the is the only fully automatic algorithm to have succeeded in flying the Berkeley helicopter. Independent evaluation also concluded that this was a significantly better controller than the hand-tweaked PD controller (which is the only other controller so far to have succeeded in keeping the helicopter in the air).

Using "potential-based reward shaping" (Ng, Harada and Russell, 1999), we have moreover trained the helicopter to fly several RC helicopter competition maneuvers. We have also succeeded in very accurately flying these difficult maneuvers. It is interesting to note that, whereas a standard way of getting a controller to fly a trajectory is by asking it to track a point that moves along this trajectory, PEGASUS learned that, in certain cases, to

fly trajectory A, it might be better to ask the helicopter to try to track some trajectory B (different from A) that reflects the varying response times of the different modes of the helicopter, so that the trajectory it actually ends up flying is a very accurate trajectory A.

We also applied PEGASUS and potential-based shaping to teaching a four-legged robot to walk. Unlike six-legged robot for which it is generally possible to design statically stable walking gaits it is very difficult to hand-design walking gaits and controllers. Indeed, the original designer of the robot parts has tried for some time to hand-design a controller, to no avail. This is also a particularly challenging problem because the robot has no sensors, and thus the robot is unable to, e.g., sense if it is falling over and try to correct for that. Instead, the robot must command the servos in a way that guarantees stability and makes progress even without closed-loop feedback. With a fairly straightforward implementation of PEGASUS, we were able to have the algorithm fully automatically learn a stable, fast, walking gait. (The robot is 10cm long, and with the learned policy walks about 10cm per second, which is quite a fast walking gait.)

4.2 Mission Planning of Multiple UAVs

We have considered the problems of decentralized optimization for multiple vehicle path planning, decentralized control structures for multiple vehicles, and decentralized estimation and control over lossy data links. Methodology is applied to multiple aircraft systems (Stanford DragonFly aircraft (2 fixed-wing, 10-foot wingspan unmanned aerial vehicles) is the target testbed).

We are examining high-level mission planning issues in the context of multiple UAVs. Each UAV acts as a mobile sensor platform which collects information about its local environment, and can communicate some of this information to its neighboring vehicles. In missions in which the goal is primarily information gathering or surveillance, it would be desirable for the vehicles to actively coordinate in order to maximize the quality of information as a whole. In missions in which the vehicles have a control objective, such as formation flight, as a stated goal, the information would be used as means of improving the tracking of the control objectives (the detection of blunders of other vehicles, for example, in order to guarantee that the formation remains collision-free).

In our model, local dynamics and local constraints are assumed decoupled from each other; the coupling in the system arises from the common (or conflicting) objectives and constraints between each subsystem. However, within the system there is no centralized decision maker. Thus each subsystem is only aware of its local model and its global constraints with no knowledge of the inner details of the other systems. We have developed a penalty-based method for decentralized optimization where coordination is achieved through a bargaining scheme. We have developed a computational testbed for rapid prototyping and analysis of coordination algorithms. The decentralized coordination network allows multiple *MATLAB* processes to run and communicate with each other using standard *TCP-IP* network protocol.

4.3 ALisp Software

We developed ALisp, an extension of Lisp that allows for the partial specification of agent programs. We have partially completed a complex testbed environment in which these

algorithms will be tested.

ALisp is a programming language for agent design that allows the user to specify partial programs which are then completed by the ALisp learning mechanism by online learning in the problem domain. Alisp consists of the Lisp language augmented with three special macros:

- `(choice <label> <form0> <form1> ... )` takes 2 or more arguments, where `<formN>` is a Lisp S-expression. The agent learns which form to execute.
- `(call <subroutine> <arg0> <arg1> ...)` calls a subroutine with its arguments and alerts the learning mechanism that a subroutine has been called.
- `(action <action-name>)` executes a "primitive" action in the MDP.

An ALisp program consists of an arbitrary Lisp program that is allowed to use these macros and obeys the constraint that all subroutines that include the choice macro (either directly, or indirectly, through nested subroutine calls) are called with the call macro.

Our work shows that the task of learning the optimal form to execute for each choice can be reduced to a SMDP. We provide several mechanisms for speeding the learning process. State abstraction allows the choices to be based on a subset of the domain variables. Since ALisp is a hierarchical programming language complete with subroutines, we can take advantage of modularity and locality by using algorithms for computing the optimal policy that follow the subroutine structure. These hierarchically structured algorithms use a 3-part method for state abstraction that allows abstraction over the reward received for each action, the reward for finishing the current subroutine, and the reward received after finishing the current subroutine. This three-part method allows considerably more state abstraction than a flat approach.

We also provide mechanisms for shaping and function approximation. Shaping allows the user to suggest desirable states and actions without committing the agent to a particular choice; the agent can recover from incorrect shaping (although it will take more time to do so). When the user provides correct shaping information, the agent learns the correct policy considerably faster. Function approximation allows the agent to learn in large or continuous domains where tabular representations for the policy and the value function are intractable.

City domain and calamity response

We have built a simulator for city driving that can support both single agent domains such as the taxi problem and multi-agent problems such as calamity response. The simulator is simple and straightfoward. City streets and blocks are represented, and vehicles have simple continuous dynamics (essentially point masses). More complex simulations for the vehicles are easily added (using functions written by Jeff Forbes that more accurately describe the motions of cars). Multiple passengers (for the taxi problem) or injured pedestrians (for the calamity response problem) can be added at any position in the world. The domain supports multiple non-controlled vehicles, so traffic is modelled and is part of the problem for both the taxi and in the calamity response domain. The domain's dynamics and representation is designed to efficiently detect collisions between vehicles.

4.4 Software for Recourse Problems

When the recourse costs of a stochastic program are defined in terms of the solution of certain partial differential equations, finding the optimal solution of such a stochastic programs becomes quite challenging. It requires marrying partial differential equations techniques with that are used to solve (simpler) stochastic programs. We are developing an algorithmic approach and the associated software when the partial differential equations are defined by a flow/transport system.

The overall objective is to develop a version the Progressive Hedging algorithm that could be used to solve stochastic optimization problems whose recourse costs are defined by a partial differential system. As test case, we consider a problem where the decision(control) consist of remediation measures affecting the underground water of a media that is heterogeneous and about which we have limited information. In this situation, there is no way to obtain an accurate description of the material properties and of the boundary and initial conditions. This means that there there no simple analytic or numerical solution to the flow/concentration equations. In turn, this implies that optimal remediation design can't be handled by 'simple' optimization techniques.

The major steps in this research have been (1) to design, and test, the numerical procedure so as to interlace the optimization steps with those involved in solving the flow/transport equations, (2) to study the relationship between the solutions (suggested decisions) with those obtained when the heterogeneous/stochastic soil is replaced by a homogenized one.

4.5 Pursuit—Evasion Scenario with Sensorweb

This section describes our work in the pursuit-evasion game (PEG) scenario, where a team of UAVs and UGVs acting as *pursuers* try to capture a group of *evaders* within a bounded but unknown environment. In cooperation with DARPA's SensorWeb project at Berkeley we have been exploring the use of the motes technology for the PEG scenario. We first discuss our work on the PEG and then discuss our work on the sensor network enhanced PEG.

PEGs

Consider a scenario in which the environment is a finite two-dimensional grid $\mathcal{X}$ with n_c square cells. $\mathcal{X}_p \subset \mathcal{X}$ ($\mathcal{X}_e \subset \mathcal{X}$) is the set of cells occupied by the n_p pursuers (n_e evaders) and w is the set of fixed obstacle locations on the ground. Ground pursuers and evaders are restricted to move to cells in which there is no other ground agent or obstacle, while aerial pursuers can move to cells in where there is no other aerial agent.

Each agent collects information about $\mathcal{X}$ at discrete time instants $t \in \mathcal{T} := \{1, 2, 3, ...\}$ and within a certain subset of the environment: *the visibility region*. We denote the visibility region of pursuer k (evader i) as $V_{p_k}(t)$ ($V_{e_i}(t)$). Each measurement $y(t), t \in \mathcal{T}$ is a triple $(v(t), e(t), o(t))$, where $v(t) \subset \mathcal{X}$ denotes the measured positions of the pursuers and $e(t) \subset \mathcal{X}$ ($\o(t) \subset \mathcal{X}$) is a set of cells where each evader (obstacle) was detected. We denoted by $\mathcal{Y}^*$ the set of all finite sequence of elements in $\mathcal{Y}$, and for each time $t \in \mathcal{T}$ denote by $Y_t \in \mathcal{Y}^*$ the sequence of measurements $\{y(1), ..., y(t)\}$ taken up to time t.

Sensor information is assumed to be perfect for the cells in which pursuers are located, but not for the other cells in the visibility region. We use a simple sensor model based on

the probability of false positives $p \in [0,1]$ and false negative $q \in [0,1]$ of an agent detecting an evader in adjacent locations. Also, we assume that the pursuers have perfect knowledge of there own locations, that is $v(t) = \mathcal{X}_p(t)$.

We assume that the pursuers are able to identify each evader separately. Therefore, pursuers keep one map for each evader and one map for the obstacles. When an evader is captured, that evader is removed from the game and its map is no longer updated. *Capture* is defined as follows: Let $x_{p_k}(t) \in v(t)$ and $x_{e_i}(t) \in e(t)$ be the estimated position of pursuer k and evader i at time t, respectively. We say that evader i is captured by pursuer k at time t if $x_{e_i}(t) \in V_{p_k}(t)$ and $d(x_{p_k}(t), x_{e_i}(t)) \leq d_m$ where $d(\cdot,\cdot)$ is a metric in $\mathcal{X}$ and d_m is a pre-specified *capture distance.* Aerial pursues can detect and share information about the positions of the evaders, but not capture an evader.

In our previous report we discussed our probabilistic map building algorithm. Since each evader is identified separately, without loss of generality, we can assume $n_e = 1$ for map building purposes. The map algorithm consists of computing the posterior probability of the evader being in cell x at time t, given the measurements Y_t taken up to time t denoted: $p_e(x|Y_t)$. Similarly, let $p_o(x|Y_t)$ be the conditional probability of having an obstacle in cell x given the measurement Y_t. There are two main steps: One, we first pool the information from the different sensors and observations to compute the posterior probability for the location of the evader. Second, given a Markov model for the evader, we can recursively update the prediction probability for the evader in the future.

We formulate the pursuit task as a multi-agent Markov decision problem. We examine two main classes of pursuit policies: greedy and global. The global policy, unfortunately, can be rather complicated to compute. This is, in part, due to the need to coordinate the agents. The greedy policy we examine is a natural heuristic in which the pursuers move towards those cells that have a high posterior probability of containing an evader. While not globally optimal the greedy algorithm has been shown, both theoretically and experimentally, to work well.

Sensor Network Enhanced PEGs

In this section we describe what a sensor network can do the PEG. There are some potential issues in the current PE framework. In particular the cameras on the UAVs and UAGs have a small range, communication among the pursuers may be difficult, unmanned vehicles are expensive, and a smart evader is difficult to catch. Thus we have started to investigate incorporating a sensor network in our PEG formulation. The benefits of a sensor network include

- large sensing coverage
- location aware sensor network provide pursuers with additional position information
- network can relay information among pursuers
- sensor network is cheap and can reduce the number of pursuers without compromising capture time
- sensibly reduce exploration of the environment
- a wide, distributed network is more difficult to compromise

The overall performance can be dramatically increased by lowering the capture time, by increasing fault tolerance, and by making the pursuer team resilient to security attacks.

Our planned testbed consists of a level field (400-2500 m^2) with 5-15 tree-like obstacles.

The pursuer team will consist of 400-1000 fixed wireless randomly placed sensor nodes with at least two modes of sensing (acoustic, magnetic). There will be 3-4 ground pursuers and 1-2 aerial pursuer. The evader's team will consist of 1-3 ground evaders with the same equipment as the ground pursuers. The ground pursuers are Pioneer UGVs and the aerial pursuers are Yamaha R50/Rmax UAVs. The communication is over wireless WaveLAN (IEEE 802.11b).

Our model is modular consisting of three main levels: sensor network, middleware services, and vehicle level sensor fusion. The middleware services will incorporate information from the sensor network including, self-location information and local time stamps and incorporate information from the pursuers including pursuer positions, evader position estimates, and global time.

There are many challenging design problems. The sensor network design involves: self-organization of the nodes, creating a communication infrastructure, self-localization, and synchronization. We are working with the DARPA SensorWeb project at Berkeley. There are also issues of,of network maintenance, robustness and security.

On the control side there are many closed loops at different levels. For the nodes we are designing algorithms that adapt to the available energy, changing physical measurements and network conditions. At the network layer we are developing energy aware algorithms for network discovery and routing. Application within the middleware level include synchronization of the agents, scheduling of actions, and localizing positions of the evaders. On the vehicles we are designing and improving our algorithms for controlling direction, stability, and probabilistic map building. Finally amongst the vehicles we are exploring multi-agent control using the competitive hidden Markov decision process methodology.

5 Collaborations

The investigators composing the MURI team interacted regularly, through meetings, joint advising of graduate students, joint teaching and collaborative research projects. In this section, we note some of the main intra-team collaborative research projects that emerged during the MURI, leading to the research described in earlier sections of the report.

- PEGASUS flies the Berkeley helicopter (Jordan, Russell, Sastry)
- Minimax classification (El Ghaoui, Jordan)
- Kernel matrix optimization (El Ghaoui, Jordan)
- Multi-agent MDPs using LP for approximate DP (Koller, VanRoy)
- MCMC for probabilistic relational logics (Russell, Koller)
- Hybrid systems analysis and control design (Tomlin, Sastry)
- Algorithms for air traffic system automation (Tomlin, Sastry, El Ghaoui)
- Variational Markov chain Monte Carlo (Jordan, Russell)
- Sparse PCA (El Ghaoui, Jordan)

6 Keynote Addresses and Invited Plenary Talks

- Keynote address at the Conference of the Association for Computational Linguistics 43rd Annual Conference (Jordan)

- Keynote address at the Conference of the American Association of Artificial Intelligence (Jordan)
- Keynote address at the International Joint Conference on Artificial Intelligence (Koller)
- Keynote address at the International Conference on Machine Learning (Jordan)
- Keynote address at the 10th International Stochastic Programming Conference (Wets)
- Keynote address at the International Conference on Machine Learning and Cybernetics (Jordan)
- Invited plenary talk at the Conference on Independent Component Analysis (Jordan)
- Invited plenary talk at the Eighth Conference on Theoretical Aspects of Rationality and Knowledge (Koller)
- Invited plenary talk at the Bar-Ilan Symposium on Foundations of Artificial Intelligence (Koller)
- Invited plenary talk at the International Conference on Cognitive and Neural Systems (Jordan)
- Invited plenary talk at the Valencia Conference on Bayesian Statistics (Jordan)
- Invited plenary talk at the IEEE Symposium 2000 on Adaptive Systems (van Roy)
- Invited plenary talk at the German/Austrian Mathematical Society Meeting (Wets)

7 Publications

E. Amir and S. Russell, "Logical Filtering." In International Joint Conference on Artificial Intelligence (IJCAI), Acapulco, Mexico, 2003.

D. Andre and S. Russell, "State Abstraction for Programmable Reinforcement Learning Agents." Advances in Neural Information Processing 14, 2001.

A. D'Aspremont, L. El Ghaoui, M. Jordan, and G. Lanckriet. "A Direct Formulation for Sparse PCA using Semidefinite Programming." SIAM Review, to appear.

F. Bach and M. I. Jordan, "Kernel Independent Component Analysis." Journal of Machine Learning Research, 3, 1–48.

F. Bach and M. I. Jordan, "Thin Junction Trees." Advances in Neural Information Processing Systems 14, 2001.

F. Bach and M. I. Jordan, "Learning Graphical Models with Mercer Kernels." Advances in Neural Information Processing Systems 16, 2003.

F. Bach and M. I. Jordan, "Tree-Dependent Component Analysis." Uncertainty in Artificial Intelligence, Proceedings of the Eighteenth Conference, 2002.

F. Bach and M. I. Jordan, "Beyond Independent Components: Trees and Clusters." Journal of Machine Learning Research, 4, 1205-1233, 2003.

F. Bach and M. I. Jordan, "Learning Graphical Models for Stationary Time Series." IEEE Transactions on Signal Processing, 52, 2189-2199, 2004.

F. Bach, G. Lanckriet, and M. Jordan, "Multiple Kernel Learning, Conic Duality, and the SMO Algorithm." International Conference on Machine Learning (ICML), New York: ACM Press, 2004.

P. Bartlett, M. Jordan, and J. McAuliffe, "Convexity, Classification and Risk Bounds," Journal of the American Statistical Association, 101, 138-156, 2006.

D. Blei, and M. Jordan. "Modeling Annotated Data." Proceedings of the 26th International Conference on Research and Development in Information Retrieval (SIGIR). New York, NY: ACM Press, 2003.

D. Blei, A. Ng, and M. Jordan. "Latent Dirichlet allocation." Journal of Machine Learning Research, 3, 993-1022, 2003.

G. Calafiore and L. El Ghaoui, "Worst-Case Maximum Likelihood Estimation in the Linear Model." Automatica, 37(4), April 2001.

J. Chou, S. Pradhan, L. El Ghaoui, and K. Ramchandran, "A Robust Optimization Solution to the Data Hiding Problem based on Spherical Codes." Proc. ICIP conference, Vancouver, Canada, October 2000.

A. Deshpande, M. N. Garofalakis, and M. I. Jordan, "Efficient Stepwise Selection in Decomposable Models." In J. Breese and D. Koller (Ed)., Uncertainty in Artificial Intelligence, Proceedings of the Seventeenth Conference, 2001.

G. Elidan, N. Lotner, N. Friedman, D. Koller, "Discovering Hidden Variables: A Structure-Based Approach." Advances in Neural Information Processing 13, 2000.

D. P. de Farias and B. Van Roy, "The Linear Programming Approach to Approximate Dynamic Programming." IEEE Transactions on Automatic Control, Vol. 51, No. 6, 2003.

D. P. de Farias and B. Van Roy, "On Constraint Sampling in the Linear Programming Approach to Approximate Dynamic Programming," Mathematics of Operations Research, Vol. 29, No. 3, August 2004, pp. 462-478.

N. de Freitas, P. Hojen-Sorensen, M. I. Jordan, and S. Russell, "Variational MCMC." In J. Breese and D. Koller (Ed)., Uncertainty in Artificial Intelligence, Proceedings of the Seventeenth Conference, 2001.

N. de Freitas, "Rao-Blackwellised Particle Filtering for Fault Diagnosis." IEEE Aerospace, 2002.

N. de Freitas and K. Barnard, "Bayesian Latent Semantic Analysis of Multimedia Databases." UBC TR 2001-15.

K. Fukumizu, F. Bach, and M. Jordan. "Dimensionality reduction for supervised learning with reproducing kernel Hilbert spaces." Journal of Machine Learning Research, 5, 73-99, 2004.

K. Fukumizu, F. Bach, and M. Jordan. "Kernel dimensionality reduction for supervised learning." Advances in Neural Information Processing 17, Cambridge, MA: MIT Press, 2004.

L. El Ghaoui and G. Calafiore. "Robust Filtering for Discrete-Time Systems with Structured Uncertainty. IEEE Transactions Automatic Control, July 2001.

L. El Ghaoui and S.I. Niculescu, "Robust Decision Problems in Engineering: An LMI Approach." In L. El Ghaoui and S.I. Niculescu, editors, Recent Advances on Linear Matrix Inequality Methods in Control, chapter 1. SIAM, 2000.

L. El Ghaoui, F. Oustry, and M. Oks, "Worst-case value-at-risk and robust portfolio allocation." Operations Research, volume 51, number 4, pp. 543-556, 2003.

C. Guestrin, M. Lagoudakis and R. Parr, "Coordinated Reinforcement Learning." In AAAI-2002 Spring Symposium on Collaborative Learning Agents, Stanford, California, March 2002.

C. Guestrin, S. Venkataraman and D. Koller, "Context Specific Multiagent Coordination and Planning with Factored MDPs." In AAAI-2002 Spring Symposium on Collaborative Learning Agents, Stanford, California, March 2002.

C. Guestrin, D. Koller and R. Parr, "Multiagent Planning with Factored MDPs." In Advances in Neural Information Processing Systems 14, Vancouver, Canada, December 2001.

C. Guestrin, D. Koller and R. Parr, "Max-norm Projections for Factored MDPs." In International Joint Conference on Artificial Intelligence (IJCAI-01), vol. 1, pp. 673 - 680, Seattle, Washington, August 2001.

G. Inalhan, D. Stipanovic, C. J. Tomlin, "Decentralized Optimization, with Application to Multiple Aircraft Coordination." 41st IEEE Conference on Decision and Control, Las Vegas, December 2002.

A. Jofre, R Rockafellar and R. Wets, "A Variational Inequality Scheme for Determining an Economic Equilibrium of Classical or Extended Type," In Variational Analysis and Applications, Springer, 2005.

H. J. Kim, R. Vidal, D. Shim, O. Shakernia, and S. Sastry, "A Hierarchical Approach to Probabilistic Pursuit-Evasion Games with Unmanned Ground and Aerial Vehicles." In Proceedings of the Conference on Decision and Control, 2001.

H. J. Kim, D. Shim, and S. Sastry, "Nonlinear Model Predictive Tracking Control for Rotorcraft-based Unmanned Aerial Vehicles." American Control Conference, 2002.

G. Lanckriet, L. El Ghaoui, C. Bhattacharyya, and M. I. Jordan, "A Robust Minimax Approach to Classification." Journal of Machine Learning Research, 3, 555-582, 2002.

G. Lanckriet, L. El Ghaoui, C. Bhattacharyya, and M. I. Jordan, "Minimax Probability Machine." Advances in Neural Information Processing Systems 14, 2001.

G. Lanckriet, L. El Ghaoui, and M. I. Jordan, "Robust Novelty Detection with Single-Class MPM." Advances in Neural Information Processing Systems 16, 2003.

G. Lanckriet, N. Cristianini, L. El Ghaoui, and M. I. Jordan, "Learning the Kernel Matrix with Semi-Definite Programming." 19th International Conference on Machine Learning, 2002.

M. Meila, and M. Jordan. "Learning with Mixtures of Trees." Journal of Machine Learning Research, 1, 1–48, 2000.

C. C. Moallemi and B. Van Roy "Distributed Optimization in Adaptive Networks," Advances in Neural Information Processing Systems 16, MIT Press, 2004.

K. Murphy and S. Russell, "Rao-Blackwellised Particle Filtering for Dynamic Bayesian Networks." In Sequential Monte Carlo Methods in Practice, A. Doucet, N. de Freitas and N.J. Gordon (eds), Springer-Verlag, 2001.

A. Y. Ng and M. I. Jordan. "PEGASUS: A Policy Search Method for Large MDPs and POMDPs." Uncertainty in Artificial Intelligence, San Mateo, CA: Morgan Kaufmann, 2000.

A. Y. Ng, D. Harada, and S. Russell, "Policy Invariance under Reward Transformations: Theory and Application to Reward Shaping." In Proc. ICML-99, Bled, Slovenia, 1999.

A. Y. Ng, M. I. Jordan, and Y. Weiss, "On Spectral Clustering: Analysis and an Algorithm." Advances in Neural Information Processing Systems 14, 2001.

A. Y. Ng and M. I. Jordan, "On Discriminative vs. Generative Classifiers: A Comparison of Logistic." Advances in Neural Information Processing Systems 14, 2001.

A. Y. Ng, A. X. Zheng, and M. I. Jordan, "Stable Algorithms for Link Analysis." Proceedings of the 24th International Conference on Research and Development in Information Retrieval (SIGIR 2001), New York, NY: ACM Press, 2001.

A. Y. Ng, A. X. Zheng, and M. I. Jordan, "Link Analysis, Eigenvectors, and Stability." International Joint Conference on Artificial Intelligence (IJCAI), 2001.

X. Nguyen, M. Wainwright, and M. Jordan, "Nonparametric Decentralized Detection using Kernel Methods." IEEE Transactions on Signal Processing, 53, 4053-4066, 2005.

M. Oishi, I. Mitchell, A. Bayen, A. Degani, C. Tomlin, "Using Formal Verification in Interface Design." IEEE CDC, March 2002.

P. Pakzad and V. Anantharam, "Estimation and Marginalization using Kikuchi Approximation Methods." Neural Computation, Vol. 17, No. 8, pp. 1836-1873, August 2005.

P. Pakzad and V. Anantharam, "A New Look at the Generalized Distributive Law." IEEE Transactions on Information Theory, Vol. 50, No. 6, pp. 1132-1155, June 2004.

P. Pakzad and V. Anantharam, "Kikuchi Approximation Method for Joint Decoding of LDPC Codes and Partial-Response Channels." IEEE Transactions on Communications. Vol. 54, No. 7, 1149-1153, 2006.

P. Pakzad and V. Anantharam, "Belief Propagation and Statistical Physics." Proceedings of the Conference on Information Sciences and Systems, Princeton University, 2002.

H. Pasula and S. Russell, "Approximate Inference for First-Order Probabilistic Languages." In Proceedings IJCAI-01, Seattle, 2001.

B. Van Roy "Performance Loss Bounds for Approximate Value Iteration with State Aggregation," Mathematics of Operations Research, Vol. 31, No. 2, pp. 234-244, 2006.

S. Russell and A. Zimdars, "Q-Decomposition for Reinforcement Learning Agents." International Conference on Machine Learning (ICML), 2003.

P. Rusmevichientong and B. Van Roy, "Decentralized Decision-Making in a Large Team with Local Information," Games and Economic Behavior, Vol. 43, No. 2, 266-295, 2003.

P. Rusmevichientong and B. Van Roy, " A Tractable POMDP for a Class of Sequencing Problems," Proceedings of the Conference on Uncertainty in Artificial Intelligence, 2001.

P. Rusmevichientong and B. Van Roy, "An Analysis of Belief Propagation on the Turbo Decoding Graph with Gaussian Densities," IEEE Transactions on Information Theory, Vol. 47, February 2001.

P. Schweitzer and A. Seidmann, "Generalized Polynomial Approximations in Markovian Decision Processes," Journal of Mathematical Analysis and Applications 110, pp. 568-582, 1985.

D. Stipanovic, R. Teo, G. Inalhan, C. Tomlin, "Decentralized Control for Interconnected Systems." Proceedings of the 41st IEEE Conference on Decision and Control, Las Vegas, December 2002.

D. Stipanovic, R. Teo, G. Inalhan, C. Tomlin, "Decentralized Overlapping Control of a Formation of Unmanned Aerial Vehicles," Automatica, Volume 40, Number 8, 1285-1296, August 2004.

S. Tatikonda and S. Mitter, "Control Under Communication Constraints," IEEE Transactions on Automatic Control, Volume 49, Issue 7, pp. 1056-1068, July 2004.

S. Tatikonda, "A Markov Decision Approach to Feedback Channel Capacity," 44th IEEE Conference on Decision and Control, December 2005.

S. Tatikonda and M. I. Jordan, "Loopy Belief Propagation and Gibbs Measures," 40th Allerton Conference on Communication, Control, and Computing, October 2002.

R. Teo, D. Stipanovic, C. Tomlin, "Cooperative Spacing Control of Multiple Vehicles over a Lossy Datalink." IEEE CDC, March 2002.

R. Teo and C. J. Tomlin, "Provably Safe Evasive Maneuvers for Blunder Situations in Closely Spaced Parallel Approaches." AIAA Guidance, Navigation, and Control Conference, January 2001.

S. Tong and D. Koller, "Active Learning for Parameter Estimation in Bayesian Networks." Advances in Neural Information Processing 13, 2000.

R. Vidal, O. Shakernia, H. J. Kim, H. Shim, and S. Sastry, "Multi-Agent Probabilistic Pursuit-Evasion Games with Unmanned Ground and Aerial Vehicles." IEEE Transactions on Robotics and Automation, 2002.

J. Yedidia, W. Freeman, and Y. Weiss, "General Belief Propagation." Advances in Neural Information Processing 13, 2000.

E. Xing, W. Wu, M. Jordan, and R. Karp. "LOGOS: A modular Bayesian Model for *de novo* Motif Detection." Journal of Bioinformatics and Computational Biology, 2, 127-154, 2004.

E. Xing, W. Wu, M. Jordan, and R. Karp and S. Russell. "A Hierarchical Bayesian Markovian Model for Motifs in Biopolymer Sequences." Advances in Neural Information Processing 16, 2003.

S. Yang, A. Kavcic, and S. Tatikonda, "Delayed Feedback Capacity of Finite-State Machine Channels: Upper Bounds on the Feedforward Capacity," 2003 IEEE International Symposium on Information Theory, July 2003.

J. S. Yedidia, W. T. Freeman, and Y. Weiss, "Bethe Free Energy, Kikuchi Approximations, and Belief Propagation Algorithms." MERL technical report, April 2000.

E. Yeo, P. Pakzad, B. Nikolic, and V. Anantharam. "VLSI Architectures for Iterative Decoders in Magnetic Recording Channels." IEEE Transactions on Magnetics, Vol. 37, No. 2, March 2001.